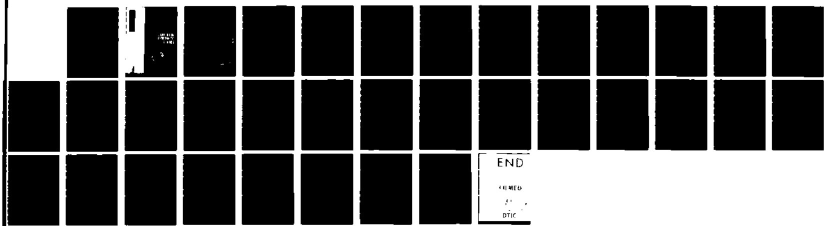
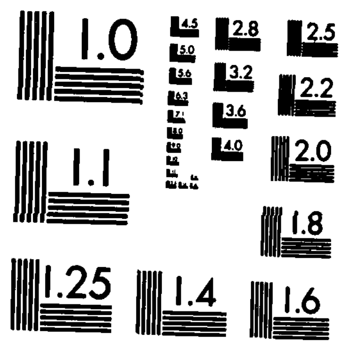
AD-A126 236 AN MDI (MINIMUM DISCRIMINATION INFORMATION) MODEL AND AN ALGORITHM FOR CO..(U) TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES A CHARNES ET AL. SEP 82 1/1

UNCLASSIFIED CCS-397-REV-2 N00014-81-C-0236 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Research Report CCS 397

(8)

AN MDI MODEL AND AN ALGORITHM FOR COMPOSITE HYPOTHESES TESTING AND ESTIMATION IN MARKETING*

by

A. Charnes
W. W. Cooper
D. B. Learner*
F. Y. Phillips*

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

Research Report CCS 397

8

AN MDI MODEL AND AN ALGORITHM FOR COMPOSITE HYPOTHESES TESTING AND ESTIMATION IN MARKETING*

by

A. Charnes
W. W. Cooper
D. B. Learner*
F. Y. Phillips*

Original: July 1981
Revised: September 1981
Second Revision: September 1982

*Market Research Corporation of America

This Research was partly supported by funds from the Market Research Corporation of America and is partly supported by Project NR 047-021, ONR Contract N00014-81-C-0236 with the Center for Cybernetic Studies, The University of Texas and ONR Contract N00014-81-C-0410 at Carnegie-Mellon University, School of Urban and Public Affairs. Reproduction in whole or part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, TX 78712
(512) 471-1821

- a -

ABSTRACT

Organizations with many different products may find it convenient to replace ad hoc statistical analyses with a uniform approach that comprehends testing and estimation. Some organizations~~, like MRCA,~~ find it imperative to move in this direction. This paper indicates how an information theoretic approach via the MDI (minimum discrimination information) statistic can be used for this purpose. Extensions to constrained versions of the MDI statistic also make it possible to test the consistency of market information with management plans or policies that can be represented in constraints formulated without reference to the data base and to estimate their impact on the market. Composite hypotheses, which are difficult to deal with by the more customary methods used in market research, can be dealt with naturally and easily via the MDI approaches. Basically MDI is more efficient than classical approaches because distribution estimation and hypothesis testing are done simultaneously. Numerical illustrations are supplied and discussed in the context of market segmentation. Developments in statistics and mathematical programming (duality) theory and methods are also briefly examined for their bearing on still further possibilities being opened for constrained MDI modeling.

KEY WORDS

Market Segmentation

Market Equilibrium

Switching Equilibrium

MDI Statistic

Composite Hypotheses

Mathematical Programming

External Constraints

1. INTRODUCTION*

This paper centers on a development of constrained information theoretic statistics with accompanying algorithms and illustrations for use in marketing. It is pointed toward composite hypothesis testing, as in a market segmentation analysis with explicitly stated constraints. Statistical testing --in multi-way contingency table analyses, for instance -- is not usually undertaken with explicitly stated arrays of constraints. The recently published book, The Information in Contingency Tables by Gokhale and Kullback, 1978, provides requisite statistical methods and rationales for such treatments. This, in turn, opens the way for contact with other parts of the management sciences where complex arrays of constraints are often used to reflect a variety of "policy" and/or "data" conditions.

Our illustrations will be effected from the data published in Ehrenberg and Goodhardt, 1974, where these data were used to show that (a) the Hendry Brand Switching Coefficient approach to market segmentation does not yield very good estimates of brand switching behavior and (b) this market -- like most markets[1/] -- is unsegmented (i.e., it consists of a single segment). They do not, however, submit their results to statistical tests.

---

*The authors are indebted to D. G. Morrison and an anonymous referee for editorial suggestions that have helped to improve this revision of earlier drafts.

1/In a private communication from G. H. Goodhardt dated September 27, 1979, he states that it has been their experience that most markets can be adequately described as being unsegmented.

We shall show how the Ehrenberg-Goodhardt results can be statistically tested relative to other alternatives by reference to their data, but we do not propose to enter into yet another discussion of the Hendry approach to market segmentation.[1/] We shall also steer clear of full-scale segmentation analysis by reference to product benefits, customer (demographic) characteristics, etc., and treat these data as only a numerical illustration for the models and methods we shall supply. The following development thus provides only a beginning but, as will be seen, it is especially well-suited for statistically testing "nested hypotheses" such as are implicit in market segmentation studies. As we shall also note (and illustrate), these constrained information theoretic models and methods possess statistical estimation as well as hypothesis testing properties that may be simultaneously exploited.

2. BACKGROUND

Marketing professionals are familiar with the use of two-way contingency tables (cross-tabulations). Formally, such two-way analyses can be extended to multi-way contingency tables, but then one confronts a variety of inadequacies in classically available statistical methods. The unsatisfactory nature of our ability to deal with large multi-way contingency tables has begun to give way before many different separate developments in the statistics and computing sciences literatures. The resulting proliferation of methods has given rise to a need for systematization accompanied by a unifying rationale based on methodological

---

[1/] See e.g., the critique in Ehrenberg and Goodhardt, 1979, of the presentation by Kalwani and Morrison, 1977.

as well as conceptual grounds. The recently published book, Discrete Multivariate Analysis, by Bishop, Fienberg and Holland, 1975, explicitly acknowledges that their effort was undertaken in response to this need.[1/] These authors use the "log-linear model" as a basis for "codifying" approaches that might be employed. The approach via log-linear models falls short of what is required, however, to supply the unifying rationale that is needed. This is supplied in Gokhale and Kullback, 1978a who use the MDI statistic for this purpose and who show how the log-linear model itself can be derived from this statistic[2/] along with a variety of other approaches. They also show how extensions beyond any of these other approaches -- e.g., including additional constraints beyond those of the contingency tables -- can be effected along these "information theoretic" lines.

We start with what, in the literature of mathematical statistics, is called the "Kullback-Leibler statistic"[3/] which is of the form

$$I(p:\pi) = \sum_{i=1}^{n} p_i \log \frac{p_i}{\pi_i} \tag{1.1}$$

where p and π are vectors with components

$$p_i, \ \pi_i \geq 0$$

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \pi_i = 1 \, . \tag{1.2}$$

1/pp. 1-3.

2/pp. 38-39.

3/After Kullback and Leibler, 1951. We have elsewhere also called this the Khinchin-Kullback-Leibler statistic -- see Charnes, Cooper and Seiford, 1978,--because of the earlier (but simpler) contribution in Khinchin, 1949.

Here the π_i may represent a set of hypothesized (constant) values which are to be tested relative to the p_i (variable) values estimated from observed data. The π_i may also represent a set of prior probabilities as in Bayesian decision theory and then the p_i become posterior probabilities determined from sample observations. Other interpretations are also possible which accord with different ways of choosing the p_i. When the components of p represent minimizing values, p*, the expression in (1.1) may be replaced by

$$(2) \qquad I(p^*:\pi) = \sum_{i=1}^{n} p_i^* \log \frac{p_i^*}{\pi_i} = \min_{p} \sum_{i=1}^{n} p_i \log \frac{p_i}{\pi_i} ,$$

and this is called the MDI (=Minimum Discrimination Information) statistic.

We may explain this name along the following lines: Consider any estimate of p in (1.1) which yields a distribution that significantly differs from the distribution as hypothesized in π. If the p distribution differs significantly from the hypothesized π distribution, might not some other estimate of p (also consistent with the data) fail to exhibit significance? The problem is resolved if p is chosen "as close as possible" to π. This is the meaning of minimum discrimination; i.e., p* provides minimum discrimination against the hypothesized π that the data, together with any other constraints, will allow.

Following the original Shannon-Wiener information theoretic formulations, the measure of information expression (1.1) was originally regarded as a probabilistic rather than a statistical concept.[1/] In his poineering book, however, S. Kullback, 1959, was able to provide a statistical foundation and to show how (1.1) and (2) could be used to unify an extremely wide variety of statistical concepts and developments.

Various statisticians, especially in the Soviet and Japanese literatures, have continued to push forward vigorously along these paths.[2/] The Japanese statistician, H. Akaike, for example, presented a very important paper at a Soviet sponsored conference[3/] in which he showed how maximum likelihood estimation (the heart of classical statistics) could be given a decision theoretic formulation asymptotically equivalent to the MDI so that the two approaches, maximum likelihood and decision theory, could be unified on one common basis. The MDI method has the unique characteristic of providing both hypothesis testing and statistical estimation regardless of the conclusion of the test. It thus also provides a decision theoretic method with unifies hypothesis testing and estimation.

Regression (sometimes of a logit or probit variety) and factor analysis are often used techniques in marketing research. Current statistical procedures usually determine the number of terms to be used on a trial and error - or exhaustive sequential -- basis. Akaike, 1973, however, shows

---

[1/] See p. 2 in Kullback, 1959.

[2/] The econometrician, H. Theil, has also pushed forward vigorously with a wide variety of publications on these topics in the American-English literatures (see e.g. Theil, 1976) and is now readying a book on information theoretic approaches to estimating statistical distributions. See Theil (in process).

[3/] See Akaike, 1973.

that the MDI approach immediately determines the number of terms to be used from the sample data by the MDI decision theoretic criterion alone.[1/]

There have been other extensions as well. Gokhale and Kullback, 1978, have extended contingency table analyses to include additional constraints. These constraints are of the form

$$\sum_j a_{ij}p_j = \theta_i, \quad i = 1, 2, \ldots, m \quad . \tag{3}$$

When the θ_i are derived from the data, as in the case of marginal totals for a contignecy table, then the constraints are said to be "internal constraints". When they are imposed on the basis of various hypothesized premises (and not by the data), as in, for instance, an assumed segmentation, then the constraints are said to be "external constraints." It is the latter which will be emphasized here if only because this provides new opportunities for dealing with managerial plans or policies that cannot be accommodated by internal constraints.

Notice the flexibility that is allowed by reference to the possibility of testing hypotheses with the π values in the functional or the θ values in the constraints. Choice of the minimizing p* in terms of (1.1), (2) and (3) makes contact with mathematical programming with its great computational power and the interpretations that are available from the underlying duality relations. These prospects, too, have now been formally effected as in Charnes and Cooper, 1974, and Charnes, Cooper and Seiford, 1978[2/] with the

[1/] Akaike also brought these methods to bear on the so-called "James-Stein paradox" wherein by the use of seemingly irrelevant data one can secure improved estimates which are not only more efficient than the mean (a maximum likelihood estimator) but which also even eliminate the mean from the admissible class of estimators in a decision theoretic sense. See Akaike, 1977, and Akaike, 1979, where the inadequacies of Bayesian approaches to this topic are also discussed. (An amusing and insightful article on the James-Stein paradox may be found in Efron and Morris, 1977, and a more general treatment of the deficiencies of maximum likelihood estimators may be found in Weiss and Wolfowitz, 1974.)

[2/] See also Charnes, Cooper and Tyssedal (forthcoming).

result that an unusually simple (unconstrained) convex programming problem is found to be the dual to minimizing (1.1) subject to (1.2) and (3)[1/], which dual is

$$(4) \qquad \max_{z} \sum_{i=0}^{n} \theta_i z_i - \sum_{j=1}^{n} \pi_j \exp\left(\sum_i z_i a_{ij}\right)$$

where $\theta_0 = 1$, $a_{0j} = 1$, $j = 1, \ldots n$ and "exp" denotes the exponential function, $e = 2.718\ldots$. See Charnes and Cooper, 1974, and Charnes, Cooper and Seiford, 1978, and observe that the choices of the components in z are not constrained.

There are other advantages that can also be secured. For example, as we have elsewhere shown -- Charnes, Cooper, Learner and Phillips, 1980c -- MDI can provide a basis for unifying a great variety of seemingly separate (and different) approaches to marketing research. Here we want to show how to bring it to bear on problems such as the composite hypothesis testing required for market segmentation testing.

3. A SEGMENTATION THEOREM AND ALGORITHM

Although information theoretic approaches can be found in the marketing literature, the example applications mainly take the form of the simpler "entropy" formula

$$(5) \qquad \sum_{i=1}^{n} p_i \log p_i$$

[1/] It is perhaps of interest to note that (1.1) provides a "proper goal functional" in the sense of Charnes and Cooper, 1977, for use in goal programming.

applied to areas like brand switching or market segmentation. Examples of such uses of the entropy concept in the marketing literature may be found in the articles by Herniter, 1973 and 1974, and by Bass, 1974[1/]. Variants may also be found in the article by Kalwani and Morrison, 1977, as well as in other articles dealing with the Hendry approach to brand switching and market segmentation analysis. See, e.g., Butler and Butler, 1971, and Hendry, 1970, 1971.

We elect to make contact with this part of the marketing literature via the data of Table 1 which Ehrenberg and Goodhardt, 1974, used to test the Hendry approach. They concluded that, on the evidence, this market does not exhibit any segmentation at all.[2/] In arriving at this conclusion, Ehrenberg and Goodhardt use a first-order analysis of the market shares. Within a Hendry analysis one must check to insure that the switching constants derived from these ratios "is applicable to the total product class as well as to the individual brands within a product class."[3/] More generally one might commence with the (relative) market share shown for each product, and then go on to consider pairs, followed by triplets, and so on, to the $2^n-1=127$ possibilities for the example shown in Table 1. The idea is to check to see whether the resulting ratios are approximately the same in order to avoid (or reduce) the danger of reaching erroneous conclusions from the (perhaps accidental) equality of lower order ratios.

Ehrenberg and Goodhardt do not undertake any of these higher order calculations, but their approach is nonetheless satisfactory by virtue of the following result

---

[1/] See, also Carter, 1979, and Haynes, Phillips and Mohrfield, 1980.

[2/] As elaborated in a private communication where, as noted on p.1 in our introduction, Ehrenberg and Goodhardt believe that most markets can be adequately described as being unsegmented.

[3/] Quoted from Kalwani and Morrison, 1977, p.470.

$$\text{If } \frac{A_1}{B_1} = \frac{A_2}{B_2} = \ldots = \frac{A_n}{B_n} \tag{6}$$

$$\text{Then } \frac{A_k}{B_k} = \frac{\sum_{s \in S} A_s}{\sum_{s \in S} B_s} \text{ for all } k \epsilon \{1, \ldots, n\} \text{ and all } S \subseteq \{1, \ldots, n\},$$

where A_j, B_k are real numbers with all $B_k \neq 0$. In other words, when the first order ratios are equal then all the higher order ratios, however they may be formed, will also be equal.

This result, which we have proved elsewhere,[1/] is employed in the first part of our proposed algorithm enabling us to follow Ehrenberg and Goodhart in restricting ourselves to first order calculations in the use of formula (7), below. We observe that (6) holds without reference to the way the A_s and B_s values are obtained and then pass on to an initial segmentation via

$$R_i = \frac{p_i - p(i,i)}{p_i\ (1 - p_i)} \tag{7}$$

where p_i = market share for brand i

$p(i,i)$ = proportion making repeat purchases of brand i.

[1/] See Charnes, Cooper, Learner and Phillips, 1980a.

Our use of Table 1 is only illustrative and so we do not make a detailed examination of the way these data were obtained or treated. Neither all rows nor all columns seem to unity--see Table 2-- so that evidently the matrix is not Markovian. Even if they did sum to unity this would not mean that the process is considered to be Markovian and the fact that Ehrenberg has been openly critical of the use of first-order Markov processes in describing consumer behavior[1/] disinclines us to treating these data in this manner. We Shall therefore interpret the p (i,i) as joint rather than conditional probabilities.

Applying (7) to the data of Table 1 we obtain the following results

$$
\text{(8)} \qquad
\begin{array}{ll}
R_1 = 0.660 & R_4 = 0.719 \\
R_2 = 0.656 & R_5 = 0.750 \\
R_3 = 0.723 & R_6 = 0.704
\end{array}
$$

Using (6) we then assert that i and j are in the same classification when R_i and R_j differ only because of sampling errors. Following Ehrenberg and Goodhardt, 1974, we have used the average of the row and column marginal values in arriving at these results. This assumes "market share equilibrium"[2/] between the two periods, a hypothesis that is subject to test along lines that we shall indicate. This, however, is not the only possible equilibrium that might be of interest for marketing purposes. For instance, we shall introduce the idea of "switching equilibrium" and show how this, too, may be tested.

[1/] See Ehrenberg, 1965, and his 1968 rejoinder to Massey and Morrison, 1968.

[2/] See Ehrenberg and Goodhardt, 1974.

Ehrenberg and Goodhardt, 1974, conclude from their analyses that this market does not segment into different classes. I.e., they conclude that there is only one "homogeneous" market in which all products compete. They do not test this hypothesis by statistical methods. As a byproduct of our use of a series of composite hypothesis testing procedures with an accompanying algorithm, we shall now show how such statistical tests may be conducted with MDI methods.

TABLE 1

The Observed Brand-Switching Percentages

for Six Brands of Breakfast Cereals

(Two successive purchases per consumer)

Product	First Purchase	Second Purchase 1	2	3	4	5	6	All Buyers
1	Corn Flakes	(23.8)	7.7	3.3	2.0	1.3	1.4	39.5
2	Weetabix	6.4	(14.3)	3.3	1.1	.8	1.1	27.0
3	Shredded Wheat	3.6	3.2	(5.4)	.8	.8	.7	14.5
4	Sugar Puffs	3.2	1.5	1.1	(3.1)	.8	.3	10.0
5	Puffed Wheat	1.7	.6	.4	.6	(1.5)	.1	4.9
6	Brand X	1.0	.4	.6	.3	.3	(1.5)	4.1
	All Buyers	39.7	27.7	14.1	7.9	5.5	5.1	100.0

4. VULNERABILITY RATIOS

These R_i values are sometimes referred to as "switching constants" -- e.g., in a Hendry analysis--but we shall not follow that usage here. We shall instead refer to them as "vulnerability ratios." Justification for this change in terminology may be presented in terms of (7) by rewriting it in the following "verbalized" form:

$$(9) \qquad R_i = \frac{1 - \dfrac{\text{repeat purchase rate for } i}{\text{brand share for } i}}{1 - \text{brand share for } i}$$

The smaller the value of this ratio, the less vulnerable is the brand in the sense that smaller R_i reflect an increasing proportion of repeat buyers in the brand's present market share. It is repeat buying on which this measure focuses. For instance, if $R_i = 0$, then the brand share for i = the repeat purchase rate for i, and <u>all</u> of this brand's customers are repeat buyers.

We keep these interpretations general and restrict our terminology accordingly. For instance, we do not assess the reasons for the observed behavior and we allow assumptions of a <u>probabilistic</u> character which may include various specialized statistical distributions used by others. As a case in point we cite Sabavala and Morrison, 1977, who use a Beta-Binomial development to obtain a "Loyalty Index" for viewing TV programs.[1/]

[1/] We are indebted to D.G. Morrison for pointing out the additional possibilities and relations derivable from Sabavala and Morrison, 1977.

In this case more special (and precise) characterizations are possible since the condition $R_i = 0$ then also implies a bipolar population in which consumers all have purchase probabilities near 0 or 1 for any such brand. Results like this suggest additional hypotheses for testing, such as situations particularized further within the situation $p_i = p_{ii}$. The latter is consistent not only with high or low p_i values but with intermediate values as well.

Another critical value of the vulnerability ratio is at $R_i = 1$ where we have

$$(10) \qquad p(i,i) = p_i^2.$$

If we interpret this in terms of probabilities, the repeat purchases are statistically independent identically distributed events. This would be the situation for a zero order process[1/] in which all consumers have the same probability, p_i, of purchasing brand i where, as we may also note, p_i may be either small or large.

For more general zero order processes involving heterogeneous populations it has been shown that $0 \leq R_i \leq 1$.[2/]

[1/] See p. 51 in Massy, Montgomery and Morrison, 1970.

[2/] We owe this result to an anonymous referee.

There are, of course, still other possibilities that may occur and so we should also allow for situations in which $R_i > 1$ as will occur, for example, if $p(i,i) = 0$, i.e., if there are no repeat purchases as in fad or fashion items. We should also like to know the implications of $R_i > 1$ if these values are to be used for monitoring possible changes in market behavior.

Now $R_i > 1$ if and only if

$$\frac{p_i - p(i,i)}{p_i (1 - p_i)} > 1 \tag{11.1}$$

so that also $R_i > 1$ if and only if

$$p(i,i) < p_i^2, \tag{11.2}$$

or the repeat purchase probability is smaller than would be obtained from statistically independent purchase behavior. Notice that the emphasis is on repeat purchases relative to other purchase behavior.

We carry this analysis a step further and rewrite (11.1) as

$$p_i - p(i,i) > p_i (1 - p_i). \tag{12.1}$$

Now

$$p_i = p(i,i) + \sum_{j \neq i} p_{ij}$$

and

$$p_i = 1 - \sum_{r \neq i} p_r$$

so that substitution in (12.1) yields

$$\sum_{j \neq i} p_{ij} > p_i \sum_{r \neq i} p_r \tag{12.2.}$$

However, $p_i \geq p\,(i,i)$ and $p_r \geq p_{ri}$ and therefore

$$(13) \qquad \sum_{j \neq i} p_{ij} > p\,(i,i) \sum_{r \neq i} p_{ri}$$

so that the probability of switching out (on the left) exceeds the probability of staying in times the probability of switching in (on the right).

The information supplied by these R_i values can be put to use in developing market strategies, with higher values of R_i being associated with more vulnerable products. Additionally these ratios may direct attention to possibilities for attracting customers from brands with R_i values exceeding one's own or of possible inroads from brands with smaller R_i values, which may have staying power because of repeat purchasing propensities.

Differences in observed R_i values may not be statistically significant, in which case the above interpretations will need to be modified. Testing for statistical significance of these and other hypotheses is a task to which we shall now turn, but our focus on this aspect of testing does no mean that we believe that other considerations such as application marketing insight, are unimportant.

5. TYPICAL FORMS

We will use specializations of (1.1) and (2) and (3) to illutstrate our procedures for testing the structure of a market in the form of a series of "nested hypotheses." We now introduce a double subscript for the discrete probability distributions with which we are concerned. The modes is[1/]

[1/] The well-known SUMT program of Fiacco and McCormick, 1968, is available for solving this kind of problem via its dual.

$$\text{Minimize} \quad I(p_{ij}:\pi_{ij}) \equiv \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{\pi_{ij}}$$

subject to

(14)

$$1 = \sum_{i,j} p_{ij}, \quad p_{ij} \geq 0, \forall\, i,j,$$

$$0 = \hat{R}_i - \hat{R}_j, \text{ etc.},$$

where

$$\hat{R}_i = \frac{1 - p(i,i)/\bar{p}_i}{1 - \bar{p}_i}$$

(15)

$$\hat{R}_j = \frac{1 - p(j,j)/\bar{p}_j}{1 - \bar{p}_j}$$

and the $\bar{p}_i$ and $\bar{p}_j$ indicate posited market share values, i.e., these values are not to be estimated by this minimization procedure. Here $p(i,i) \equiv p_{ii}$. These values are to be estimated along with the p_{ij} which represent proportions switching between products i and j over the two purchase occasions.

The estimated $\hat{R}_i = \hat{R}_j$ represent vulnerability ratios hypothesized to accord with "nestings" indicated by arrangements like

$$0 = \hat{R}_i - \hat{R}_j$$

(16)

$$0 = \hat{R}_j - \hat{R}_k$$

when products i, j and k are hypothesized to be in the same vulnerability class.

To simplify notation we omit the circumflexes on the R_i, R_j, R_k and designate optimal estimates by

(17)

$$R_i^* = \frac{1 - p_{ii}^*/\bar{p}_i}{1 - \bar{p}_i}$$

$$R_j^* = \frac{1 - p_{jj}^*/\bar{p}_j}{1 - \bar{p}_j}$$

The resulting

$$(18) \qquad I^* = I\,(p^*_{ij}\,,\,\pi_{ij})$$

may then be used to test the hypothesis that the minimizing p^*_{ij} choices do not deviate from the distribution associated with the π_{ij}. The π_{ij} represent "switching proportions" hypothesized in our case to accord with the data represented in Table 2,

TABLE 2

i \ j	1	2	3	4	5	6
1	.238	.077	.033	.020	.013	.014
2	.064	.143	.033	.011	.008	.011
3	.036	.032	.054	.008	.008	.007
4	.003	.015	.011	.031	.008	.003
5	.017	.006	.004	.006	.015	.001
6	.010	.004	.006	.003	.003	.015

where the π_{ij} data of Table 2 are the Table 1 data expressed as fractions. The p^*_{ij} values are to be secured by solving (14) using <u>all</u> of the data (off-diagonal as well as diagonal) in Table 2.

A characterization which provides access to the statistical properties as developed in Kullback, 1959, for the MDI statistic may be obtained from the widely shared belief that brand switching is proportional to brand share[1/]

[1/] See Ehrenberg and Goodhardt, 1974, where this is stated on p. 232 as "the probability of switching from i to j depends only on the purchase probabilities p_i and p_j (which are equivalent to market shares)."

and with R_i^* regarded as a switching constant, this is reflected in (7).

Although this empirical "law" has been regarded as implying that one has switching equilibrium in a zero order process, there are other less restrictive assumptions which can guarantee such observations empirically. For example, suppose that the underlying distribution which yields these switches is multinomial in character with probabilities p_i, $i = 1, 2, \ldots, n$. It then follows from the multivariate form of the central limit theorem that the observations tend to the multi-normal distribution whose covariance matrix is symmetric with off-diagonal terms proportional to the product of the corresponding brand share.[1/] Thus, for large samples, one should then expect to find "brand switching proportional to the product of brand shares." Conversely, a finding of statistically significant deviations will result in rejection of the empirical "law" that switching is proportional to brand share.

6. ALGORITHM AND TESTING PROCEDURE

The results provided in (6) allowed us to reduce the number of brand combinations to be considered in a market segmentation analysis. This is important also when we come to effecting our statistical tests since one might (in principle) have to consider a great number of possible $\hat{R}_i$, $\hat{R}_j$ pairs to test in (15).

1/For a more detailed discussion with accompanying further references, see Charnes, Cooper, Learner and Phillips, 1980a.

We therefore suggest the following algorithm which is based on (7) in section 3:

> Order the R_i values from smallest to greatest as $R_{(1)}, R_{(2)}, \ldots, R_{(n)}$, as in Table 3 below. See (8). By MDI, test the hypothesis that $R_{(1)}$ and $R_{(2)}$ are equal at a specified level of significance. If the resulting value of the MDI statistic rejects the hypothesized equality then segment $R_{(1)}$ from $R_{(2)}$. Then begin with $R_{(2)}$ as a smallest R_i to test for a further segment apart from $R_{(3)}$. If, instead, the test accepts the hypothesized equality of $R_{(1)}$ and $R_{(2)}$, add the additional condition to $R_{(2)} = R_{(3)}$. Test with the MDI statistic. If the test accepts the hypothesized equality of $R_{(2)}$ and $R_{(3)}$ go on to new higher R_i by adding the new pertinent additional equality conditions.

To illustrate our suggested procedures for using these MDI values, we now refer to the following R_i values from the results portrayed in (8), but arranged as indicated by our algorithm.

TABLE 3

i:	2	1	6	4	3	5
R_i:	.656	.660	.704	.719	.723	.750

Applying this algorithm we test the hypotheses indicated by the nestings (from top to bottom) under the column labelled "Null Hypothesis" in Table 4. With this formulation, Ehrenberg and Goodhardt apparently are correct, or at least their "one homogeneous market" conclusion cannot be rejected on the basis of these data for all sample sizes, $N \leq 6{,}500$ and significance level $\alpha = 0.95$.

TABLE 4

Problem	Null Hypothesis	I*	d.f.#	$N \leq 6500$ $\alpha = 0.95$
1	$R_2 = R_1$	Negligible	1	accept
2	$R_2 = R_1 = R_6$	1.24×10^{-4}	2	accept
3	$R_2 = R_1 = R_6 = R_4$	3.60×10^{-4}	3	accept
4	$R_2 = R_1 = R_6 = R_4 = R_3$	5.94×10^{-4}	4	accept
5	$R_2 = R_1 = R_6 = R_4 = R_3 = R_5$	8.54×10^{-4}	5	accept

#The degrees of freedom are generally equal to the number of linearly independent constraints--with a rigorous derivation relating them to the number of parameters in the dual system. See. e.g., pp. 38-39 and 181 ff. in Gokhale and Kullback, 1978a. See also Gokhale and Kullback 1978b.

In this case we have not rejected the hypothesis of no market segments. In the opposite case we would simply push to a next level of analysis by similarly testing for equality of the vulnerability ratio within each segment using multiple subscripted R_is. Instead of testing such hypothesized segments one at a time, we could simultaneously test for equality (e.g., equality with different switching constants) in each hypothesized segment while refraining from imposing similar constraints on other parts of the market.

We can also test for market homogeneity in other ways. For instance we can test for market homogeneity based on requiring the estimated p_{ij} to conform to the conditions for switching equilibrium. This requires the adjunction of additional constraints as in the following:

$$\min I\ (p:\pi)$$

subject to

$$(19)\qquad \begin{aligned} &p_{ij} = p_{ji}, \quad \forall i \neq j \\ &\sum_{i=1}^{6} \sum_{j=1}^{6} p_{ij} = 1 \\ &p_{ij} \geq 0 \quad \forall i,j \ . \end{aligned}$$

The resulting estimates are portrayed in Table 5 with $I^* = 0.007$ and 15 degrees of freedom. The hypothesis of "switching equilibrium" is <u>not</u> rejected for all $N \leq 1430$ at $\alpha = 0.95$.[1/]

TABLE 5

p^*_{ij} Switching Equilibrium Values , $i \neq j$

i \ j	1	2	3	4	5	6
1	.240	.071	.035	.025	.015	.012
2	.071	.144	.033	.013	.007	.007
3	.035	.033	.054	.009	.006	.007
4	.025	.013	.009	.031	.007	.003
5	.015	.007	.006	.007	.015	.002
6	.012	.007	.007	.003	.002	.015

[1/] Because we are using these data only for illustration we do not adjust these α values in these successive tests.

Apparently the market portrayed in Table 1 is consistent with the hypothesized switching equilibrium and the p_{ij}^* values portrayed in Table 5 represent "best" estimates of the switching probabilities in the sense that they are "as close as it is possible to get" to the π_{ij} values in Table 2 while retaining the symmetry conditions for switching equilibrium as reflected in (19).

Turning next to the question of market-share equilibrium we now formulate our model for purposes of testing this hypothesis by incorporating these conditions explicitly as follows:[1/]

$$\min \; I\,(p:\pi)$$

subject to

$$\sum_{j=1}^{6} p_{ij} = \bar{p}_i$$

$$\sum_{i=1}^{6} p_{ij} = \bar{p}_j$$

$$(17) \qquad \hat{R}_i = \hat{R}_j$$

$$\vdots$$

$$\hat{R}_k = \hat{R}_\ell$$

$$p_{ij} \geq 0, \qquad \forall\, i,j$$

where $\bar{p}_i$ and $\bar{p}_j$ represent the average of the rim values from Table 2 with $\bar{p}_i = \bar{p}_j$ whenever $i=j$ and $\sum_{i=1}^{6} \bar{p}_i = \sum_{j-1}^{6} \bar{p}_j = 1$. In other words, we formalize the assumption of an hypothesized market share equilibrium and then repeat the

[1/]Compare with (14) and (15). Although Ehrenberg and Goodhardt are clear in their discussion, no such tests were conducted by them -- possibly because classical statistical mechanisms are not designed to deal with such explicitly constrained models involving "external constraints" in rather complex arrays.

same tests as in (14) to ascertain whether the conclusion of a single homogeneous market is correct and obtain the results shown in Table 6.

TABLE 6

MARKET SEGMENTATION WITH
MARKET SHARE EQUILIBRIUM

Problem	Null Hypothesis	I*	d.f.	N=3750 α=0.95
1	$R_1=R_2$	2.78×10^{-3}	12	accept
2	$R_1=R_2=R_6$	3.02×10^{-3}	13	reject
3	$R_1=R_2$; $R_6=R_4$	2.88×10^{-3}	13	accept
4	$R_1=R_2$; $R_6=R_4=R_3$	2.89×10^{-3}	14	accept
5	$R_1=R_2$; $R_6=R_4=R_3=R_5$	2.99×10^{-3}	15	accept

For $N < 3{,}750$ the Ehrenberg-Goodhardt conclusion of a single homogeneous market continues to be maintained as before. For $N = 3{,}750$ the hypothesis of market homogeneity would be rejected since at this sample size (and above) the market segments into two. Thus at these sample sizes the market stationarity which Ehrenberg and Goodhardt assumed in their averaging is not sustained.

We also provide the p^*_{ij} estimates corresponding to the hypothesis of problem 5 in table 6. These are presented in Table 7. Note that these values are consistent with the indicated segmentation in $N = 3{,}750$. At this sample size, the results in Tables 4 and 6 are consistent and at $N \leq 1{,}430$ they are also consistent with the concept of switching equilibrium. At sample sizes larger then $N = 1{,}430$, however, the hypothesis of switching equilibrium is rejected while the hypothesis of market - share equilibrium

continues to be maintained for 1,430 $\leq$ N $\leq$ 3,750. Evidently, the two concepts are not the same. Thus, even with an hypothesized and validated market-share equilibrium the vulnerability ratio continues to provide valuable information by the way it summarizes the net switches in and out and by the way it designates the more vulnerable products by reference to repeat buying relative to market share.[1/]

TABLE 7

ESTIMATES OF SWITCHING/REPEAT PROBABILITIES IN A TWO-SEGMENT MARKET UNDER MARKET-SHARE EQUILIBRIUM

i \ j	1	2	3	4	5	6
1	.239	.076	.033	.024	.012	.012
2	.065	.143	.034	.014	.007	.010
3	.035	.031	.054	.010	.007	.006
4	.028	.013	.010	.031	.006	.002
5	.017	.006	.004	.007	.016	.001
6	.012	.005	.007	.004	.003	.014

[1/] See the discussion in Section 4.

Supposing that the sample size is adequate, the estimates in Table 7 provide further insight into the Table 6 results. Evidently the switching probabilities between brands 1 and 2 are high relative to the rest of the market. Moreover, the repeat-purchase probabilities for each of brands 1 and 2 are also relatively high.

Presumably this finding would next be expanded in terms of further analyses such as whether this segment exhibited brand-primary or form-primary characteristics, etc.[1/] Our interpretation of the R_i values as "vulnerability ratios" can also be used to yield conclusions of possible managerial significance which we can illustrate with the results in Tables 6 and 7. Thus, via the results in Table 6, the marketing managers of brands 1 and 2 might be advised to turn their attention to the segment identified via $R_6 = R_4 = R_3 = R_5$. As is seen in Table 7, the repeat purchase probabilities in this segment are all relatively low. The switching probabilities between the brands in this segment are also relatively low and the fact that they are all classified together by our statistical tests means that the marketing managers of brands 1 and 2 should orient their strategy considerations to the whole (equally vulnerable) segment. To state this last point somewhat differently, if our statistical tests had identified a further decomposition then the marketing managers of brands 1 and 2 would be better advised to consider the different vulnerability values and the strategies suitable to taking advantage of them.

[1/] A third "mixed" characterization is also admitted as in Robinson, Vanhonacker and Bass, 1980.

CONCLUDING REMARKS

Our illustrations have been confined to the case of external constraints. Examples using only internal constraints (which are necessarily consistent with the observations)[1/] could also be employed. Combinations of internal and external constraints are possible, of course, but these involve more recondite analyses since a general theory for their treatment is not yet available.

Our illustrations also involve only linear equality constraints where, again, the theory is fully developed and readily available. More general non-linear relations and even inequality constraints are only now beginning to be addressed.[2/]

Even while allowing for these limitations, a great variety of other possibilities are also present besides those illustrated in this article, and the accompanying algorithm and models we have provided should help to provide at least a start for the use of such MDI methods in marketing. This includes possibilities for further testing of sharpened hypotheses when one is willing to make more specialized assumptions about the statistical distributions and/or (as in Sabavala and Morrison, 1977) the underlying probability models governing the behavior of individual consumers. It also opens contacts with mathematical programming (and its duality relations) which have proved of great value in other types of modelling not only for the computational power that is provided but also for a great variety of managerial uses and interpretations including those of a policy evaluation variety. The way is opened for using such models with their explicitly

---

1/See p. 181 in Gokhale and Kullback, 1978a.

2/See Charnes, Cooper and Tyssedal (forthcoming) for a recent extension of the duality relations described in Charnes, Cooper and Seiford, 1978, to comprehend situations involving inequality constraints.

formulated constraints to reflect a variety of managerial policy conditions (e.g., on market share, etc.) which can then be tested and evaluated. Proposed policies and marketing activities as reflected in external constraints can then be related to costs and benefits flowing from their imposition or accomplishment. The ultimate managerial benefit is that approaches along these lines can serve to direct management to activities that ought to be considered and to provide a way of evaluating various mixtures of policies relative to their possibilities for realization in the market.

REFERENCES

1. Akaike, H., "Information Theory and an Extension of the Maximum Likelihood Principle." 2nd International Symposium on Information Theory, (B.N. Petrov and F. Csaki, eds.) Akademiai Kiaido, Budapest, (1973) 267-281.

2. ________, "An Extension of the Method of Maximum Likelihood and the Stein's Problem." Institue of Statistical Mathematics 29, Part A, (1977) 153-164.

3. ________, "A New Look at the Bayes Procedure." Biometrika 65, No. 1, (1978) 53-59.

4. Bass, F. M., "The Theory of Stochastic Brand Preference and Brand Switching." Journal of Marketing Research, 11, Feb., 1974, pp. 1-20.

5. Bishop, Y. M. M., S. Fienberg, and P. W. Holland, Discrete Multivariate Analysis, Cambridge, MA: MIT Press, (1975).

6. Blattberg, R. C. and S. K. Sen, "Market Segmantation and Stochastic Brand Choice Models." Journal of Marketing Research 13, Feb. 1976, pp. 34-45.

7. Brockett, P. L., A. Charnes and W. W. Cooper, "M.D.I. Estimation via Unconstrained Convex Programming." Research Report CCS 326 (Austin, TX: University of Texas, Center for Cybernetic Studies, 1978).

8. Butler, D. and B. Butler, Jr. Hendro-Dynamics: Fundamental Laws of Consumer Dynamics, (Cambridge, MA: The Hendry Corporation, Chapter 1, 1970, and Chapter 2, 1971).

9. Carter, J. C., An Entropic-Based Partitioning Approach to the Analysis of Competition, Ph.D Thesis, Columbia University, 1975 (Ann Arbor, MI: University Microfilms International, 1979).

10. Charnes, A. and W. W. Cooper, "Constrained Kullback-Leibler Estimation, Generalized Cobb-Douglas Balance and Unconstrained Convex Programming." Rendiconti di Academia Nazionale dei Lincei, Series VIII, Vol. 56, April, 1974.

11. ________ and ________, "Goal Programming and Multiple Objective Optimizations; Part I." European Journal of Operational Research 1, No. 1, Jan. 1977, pp. 39-54.

12. ________ and ________, Management Models and Industrial Applications of Linear Programming, (New York: John Wiley & Sons, Inc., 1961).

13. __________ and __________, and D.B. Learner, "Constrained Information Theoretic Characterizations in Consumer Purchase Behavior," Journal of the Operational Research Society, No. 9, 1978, pp. 832-842.

14. __________, __________, __________, and F.Y. Phillips, "A Theorem and Approach to Market Segmentation," in R.P. Leone, ed., Proceedings -- Market Measurement and Analysis, published by TIMS College of Marketing and TIMS, 1980 a.

15. __________, __________, __________, __________, "An MDI Procedure for Vulnerability Segmentation Tests," (Austin, TX: The University of Texas, Center for Cybernetic Studies, Research Report CCS 376, Sept. 1980 b).

16. __________, __________, __________, __________, "The MDI Method as a Generalization of Logit, Probit and Hendry Models in Marketing," Market Research Corporation of America PDA-RP#70, 1980c.

17. __________, __________, and L. Seiford, "Extremal Principles and Optimization Dualities for Khinchin-Kullback-Leibler Estimation," Mathematisch Operationsforschungund Statistik, Series Optimization, Vol. 9, No. 1 (1978)

18. __________, __________, and J. Tyssedal, "Khinchin-Kullback-Leibler Estimation with Inequality Constraints," Mathematisch Operationsforschungund Statistik, Series Optimization,(forthcoming).

19. Efron, B., and C. Morris, "Stein's Paradox in Statistics,", Scientific American, 236, No. 5, May 1977, pp. 119-127.

20. Ehrenberg, A.S.C, "An Appraisal of Markov Brand-Switching Models," Journal of Marketing Research 2, 1965, pp. 347-373. See also A.S.C. Ehrenberg "On Clarifying M and M," Journal of Marketing Research 5, 1968, pp. 228-229.

21. __________, and G.J. Goodhardt, "The Hendry Brand Switching Coefficient," ADMAP, August 1974, pp. 232-238. Also available from London, ASKE Research Ltd.

22. __________, __________, "The Switching Constant," Management Science 25, No. 7, July 1979, pp. 703-705.

23. Fiacco, A.J. and G.P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques (New York: John Wiley & Sons, Inc. 1968).

24. Gokhale, D.V., and S. Kullback, The Information in Contingency Tables, (New York: Marcel Dekker, New York, 1978a).

25. __________, __________, "The Minimum Discrimination Information Approach in Analyzing Categorical Data," Communication in Statistics A. Theory and Methods, A7(10), 987-1005, 1978b.

26. Haynes, K.E., F.Y. Phillips and J.W. Mohrfeld, "The Entropies: Some Roots of Ambiguity," Socio-Economic Planning Sciences, May (1980).

27. Hendry Corporation, Hendrodynamics: Fundamental Laws of Consumer Dynamics, Chapter 1 (1970), Chapter 2 (1971)

28. Herniter, J., "An Entropy Model of Brand Purchase Behavior," Journal of Marketing Research, Vol. X, Nov. 1973.

29 __________, "A Comparison of the Entropy Model and the Hendry Model," Journal of Marketing Research, Vol. XI, Feb. (1974).

30. Kalwani, M.U., "The Entropy Concept and the Hendry Partitioning Approach," (Cambridge, MA: Sloan School of Management, MIT, WP#1072-79, June 1979).

31. __________, and D.G. Morrison, "A Parsimonious Description of the Hendry System," Management Science 23, No. 5, Jan. (1977), pp. 467-477.

32. Khinchin, A.I. Mathematical Foundations of Statistical Mechanics (New York: Dover Publications, 1949).

33. Kullback, S, Information Theory and Statistics, John Wiley, New York (1959)

34 __________, and R.A. Leibler, "On Information and Sufficiency," Annals of Mathematical Statistics 22, pp. 79-86 (1951).

35. Learner, D.B., and F.Y. Phillips, "Controllable Forecasting in Marketing," Presented at ORSA/TIMs Conference on Market Measurement and Analysis, Stanford CA, March 1979).

36. Massy, W.F., D.B. Montgomery and D.G. Morrison, Stochastic Models of Buying Behavior (Cambridge: The MIT press, 1970).

37. __________, and D.G. Morrison, "Comments on Ehrenberg's Appraisal of Brand-Switching Models," Journal of Marketing Research 5, 1968, pp. 225-228.

38. Phillips, F.Y., Some Information Theoretic Mehtods for Management Analysis in Marketing and Resources, Ph.D. Dissertation, University of Texas at Austin, 1978).

39. __________, "A Guide to MDI Statistics for Planning and Manage-Model Building," Institute for Constructive Capitalism Technical Series #2, The University of Texas at Austin (1980).

40. Rubinson, J.R., W.R. Vanhonacker and F.M. Bass, "A Note on 'A Parsimonious Description of the Hendry System'," Management Science 26, 1980, pp. 215-226.

41. Sabavala, D.J. and D.G. Morrison, "A Model of TV Show Loyalty," Journal of Advertising Research 17, No. 6, Dec. 1977, pp. 35-43.

42. Theil, H., The Maximum Entropy Distribution (in process).

43. ________, Statistical Decomposition Analysis (New York: John Wiley & Sons, Inc. 1976).

44. Weiss, L. and J. Wolfowitz, Maximum Probability and Related Topics, (Berlin: Springer-Verlag, 1974).

4-8

DTI